



**PARIS SCHOOL OF ECONOMICS**  
ÉCOLE D'ÉCONOMIE DE PARIS

**WORKING PAPER N° 2013 – 12**

**How Do You Feel?  
The Effect of the New Cooperative Medical Scheme in China**

**Carine Milcent  
Binzhen Wu**

**JEL Codes: D82, H52, I18**

**Keywords: Subjective health; Rural China; Health insurance; New Cooperative Medical Scheme**



**PARIS-JOURDAN SCIENCES ÉCONOMIQUES**

48, Bd JOURDAN – E.N.S. – 75014 PARIS  
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10  
[www.pse.ens.fr](http://www.pse.ens.fr)

# How Do You Feel?

## The Effect of the New Cooperative Medical Scheme in China\*

Carine Milcent and Binzhen Wu

May 28, 2013

### Abstract

During the 2003-2006 period, subjective health status in Chinese rural areas improved. We used a unique household longitudinal survey to analyze how the introduction of a public insurance system has contributed to the change. This program is based on a doubly voluntary process: counties decide to launch, then households decide to subscribe. We disentangle two channels of influence of the insurance: the insurance effect of the coverage and a general equilibrium effect on all residents in the insurance-adopting counties. The empirical findings include, first, a positive extensive margin: individuals feel better about their health status when covered by the NCMS. However, there is no intensive margin: an individual's self-assessed health status does not improve with the number of years enrolled in the program. Second, we find a positive general equilibrium effect of introducing the NCMS program on non-participants in NCMS counties. This effect accumulates over time.

Keywords: Subjective health, Rural China, Health insurance, New Cooperative Medical Scheme.

JEL Classification: D82, H52, I18.

---

\*We would like to thank Chong-en Bai for his helpful support and for giving us access to the data and the seminar participants of Health Economics at TsingHua University, Beijing, China. We are also grateful to Thomas Serrurier for helpful comments.

# 1 Introduction

In recent years, the self-reported health of the Chinese rural population has worsened. The percentage who declare themselves to be in bad health rose by about 15 percentage points between 1991 and 2004 (Milcent, 2012). Over the same period, there has been a drastic decline in the utilization of health care (Bloom and Gu, 1997; Hesketh and Zhu, 1997; Gao et al., 2001). A common explanation of these two phenomena is the rise in the price of health care and the absence of a health insurance market. One of the major causes of impoverishment in rural China is now the financial hardship caused by ill health (Liu, 2006). In 2003, 80% of rural residents in China did not have any health insurance (Wagstaff et al., 2009). In the same year, a new health-insurance system, the New Cooperative Medical Scheme (NCMS), was introduced in rural areas. By 2009, more than 94% of eligible rural residents (833 million people) were covered by the NCMS (China Statistical Yearbook 2010). This paper uses a unique household longitudinal survey which allows us to distinguish both *i)* the effect of the health-insurance market on self-assessed health (SAH), and *ii)* the effect of insurance coverage on SAH.

Individuals do of course have some idea about their own health status, but this knowledge depends on their medical know-how and the frequency with which they go to the hospital or to the doctor. The need of individuals for healthcare is determined by their objective health, which is only partially known to them. The perception that individuals have of their own health does not determine their need for healthcare but it determines their demand for healthcare access. The role of public insurance is to improve the affordability of healthcare, and allow individuals to fulfill their demand for healthcare access without, or at least with less regard to, their budget constraints. Hence, public insurance is a political tool which responds to healthcare demand based on individuals' subjective health. The NCMS program offers a quasi-natural experiment which allows us to address a novel question: How does the introduction of a health insurance program change individuals' views of their subjective health?

Moreover, insurance has the effect of modifying the individual's perceived health, via greater use of healthcare. On the one hand, we expect the individual's health to improve, via the greater use of healthcare associated with looser budget constraints; on the other hand, more frequent healthcare use should also give individuals more accurate knowledge of their objective health status, which will also affect their perceived health. Moreover, the presence of health insurance may also affect the supply of healthcare. As individuals become less budget-constrained, the

supply of healthcare may become both better-quality and more expensive. As such, the mere presence of health insurance may have an impact on the healthcare demand of those who are not covered by the insurance. The budget constraints of the latter may be affected because of these changes in the supply of healthcare; a resulting effect on their subjective health may then be expected. Last, as the use of healthcare rises, it is likely that a greater amount of information about health will circulate between individuals. These assumptions will be tested in this paper.

Although the NCMS has been widely seen as one of the most important reforms in rural China since the economic reforms of the 1980s, the systematic empirical evidence on this issue is very limited, with a noticeable exception in Lei and Lin's (2009) finding of no significant improvement in self-reported health status or sickness or injury in the past four weeks after families acquired insurance coverage. Similar to the studies on public health insurance programs in other countries, the few papers that have evaluated the NCMS program have all focused on its effects on healthcare utilization or the healthcare spending of households. However, it is not clear whether more (less) healthcare utilization means better (worse) self-assessment of health status,<sup>1</sup> not to mention that evidence in the literature on the effect of the NCMS on healthcare utilization is actually mixed (Wagstaff et al., 2009; Lei and Lin, 2009; Bai and Wu, 2010). Despite the fact that the main purpose of the NCMS is to improve the affordability of healthcare services, empirical evaluations have not found solid evidence of a decrease in out-of-pocket health expenditures (per visit) owing to the NCMS (Wagstaff et al., 2009; Lei and Lin, 2009; Bai and Wu, 2010). It has been suggested that suppliers may have exploited the benefit by increasing the price of healthcare services or increasing the number of visits or unnecessary tests or procedures for each patient, resulting in an increase in total healthcare expenses. This increase can also affect both people's health status and their assessment of it. For example, a decrease in the affordability of healthcare for the uninsured may have a negative effect on their views of their health status. Moreover, the NCMS program affects not only insured families, but also non-participants in the counties that have introduced the program (labeled as NCMS counties). We call these effects general equilibrium effects. These generally do not directly result from better healthcare. There are also other general equilibrium effects, including a reinforcement of awareness of health risks when the government tries to convince people to participate to the program, the spillover of health-related knowledge,

---

<sup>1</sup>On one hand, more healthcare utilization can lead to better objective and hence subjective health status; on the other hand, more utilization of the healthcare may lead to more awareness or anxiety of latent diseases (Koszegi, 2001).

and peer-comparison effects when exogenous shocks affect the health status of individuals in the community.

Due to the fact that the NCMS has been rolled out sequentially in a series of counties,<sup>2</sup> and that counties and households decide whether to participate voluntarily, three types of households can be distinguished in each year: the insured, non-participants in villages within a NCMS county, and the non-exposed in villages within counties that have not launched the program (labeled as non-NCMS counties). Three effects of the program can be identified after addressing potential selection biases: the insurance effect, which works only on the insured, by comparing the insured with non-participants; the gross effect of the program, by comparing the insured with the non-exposed; and the spillover effect, by comparing non-participants with the non-exposed. Previous studies have focused on either the insurance effect or the gross effect of the program: our paper distinguishes between them.

This paper draws upon a unique dataset that combines a longitudinal survey carried out in rural China (National Fixed-Point Survey, or NFS) from 2003 to 2006 and a household survey on the NCMS that we conducted with a subsample of the 2006 round of the RFPS. The way the NCMS was introduced provides a quasi-natural experiment which can be used to examine how this public insurance program changed people's self-assessment of health. Nevertheless, because county enrollment is not completely random and participation is voluntary, the main challenge of the study is selection bias among both counties/villages and households. To deal with selection bias, we combined a matching method with the difference-in-difference framework. We found, first of all, that individuals feel better about their health status when they are covered by the NCMS. However, there is no additional gain in self-assessed health status with additional years of enrolment in the program. Second, we find a positive general equilibrium effect of the introduction of the NCMS program on non-participants in the NCMS county, which accumulates over time. Third, besides the insurance program, household composition (family size, number of young and old members) and socio-economic status (income, occupation, and education) also have strong effects on self-evaluated health status.

The paper is organized as follows: Section 2 summarizes the literature and describes the Chinese context; Section 3 describes the data and gives preliminary statistics; Section 4 explains the econometric strategy; Section 5 presents the results, and Section 6 concludes.

---

<sup>2</sup>A county is composed of villages

## 2 Literature Review and Chinese Context

### 2.1 The New Rural Community Medical System (NCMS)

Along with the economic reforms of the 1980s came the dissolution of the rural Cooperative Medical System, and illness emerged as a leading cause of poverty in rural China. The high cost of healthcare has deterred many families from obtaining necessary healthcare. In response, the Chinese government started pilot programs of the New Cooperative Medical System in 2003. NCMS programs have several key features: 1) the program is targeted at rural residents;<sup>3</sup> 2) participation is voluntary but must be at the household level; 3) participants are required to pay flat-rate premiums, but their contributions are heavily subsidized by governments; 4) the programs mainly reimburse large expenses so as to ease the economic burden resulting from catastrophic disease and alleviate illness-caused poverty; 5) the programs are operated at the county level rather than the township or village level. The primary goal of the NCMS is to reduce impoverishment resulting from illness and improve the affordability of healthcare (Central Committee of CPC, 2002). Local governments have been granted autonomy to design, implement and supervise the programs.

Participation in NCMS insurance is voluntary but the unit of participation is the household. There is reason to assume that household subscriptions are based on the average health status of members of the household, but not that they are based on every member of the household having a poor health status: the system thus reduces adverse selection bias. However, although rural inhabitants are required to participate as household units in order to reduce adverse selection, the elementary conclusion drawn from previous studies is that the system cannot prevent adverse selection (Wagstaff *et alii*, 2007). In this paper, we control for household selection bias.

While local governments have some discretion over the level of financing of the program, the standard in 2003 was for each participating household to pay at least 10 RMB (about €1 and about \$1.2) for each household member every year, with the local government providing more than 10 RMB for each person per year. The central government would also match with 10 RMB per year for each beneficiary living in the central and western provinces. As of 2006, while individual contributions generally remained at the existing level, subsidies from local and central government increased (Wagstaff *et al* 2007). NCMS insurance mainly provides financial risk protection to

---

<sup>3</sup>The program will also be offered in urban districts and county-level cities that include rural residents.

patients with catastrophic health problems. Many services—particularly outpatient care—are not covered, deductibles are high, ceilings are low, and coinsurance rates are high.

There are more than 2800 rural counties in China. The NCMS pilot program began in 310 rural counties in 2003. It expanded to 617 counties in 2005, 1451 counties in 2006—i.e., 50.7% of the total number of counties—and started to spread across the country in 2007. By the end of June 2007, the program had expanded to cover 84.9% of all rural counties and 82.8% of all rural residents. It has covered the whole rural population since 2010. Provincial and county governments retain considerable discretion over the details of the pilots, including the placement of the pilot program.

In fact, NCMS pilot counties were not randomly selected. Instead, a complex set of criteria, including local interest and capacity, level of economic development, and the status of the delivery system were considered. The voluntary nature of NCMS raises concerns about an adverse selection that may represent a serious threat to the financial sustainability of the NCMS. In this paper, we control for the county selection bias: in other words, we control for the effect of living in a county selected to have an NCMS on self-assessed health (SAH).

Along with the NCMS, local governments provide some supporting policies, such as improvements in rural healthcare (delivery) networks and health service provision, the strengthening of pharmaceutical governance, and supply chain construction. These improve the quality and delivery of healthcare service and also benefit families who choose not to participate the NCMS. Therefore, the adoption of the NCMS program by the county may have a positive effect on healthcare access for both the NCMS-insured and the NCMS-uninsured.

NCMS insurance programs make healthcare services accessible to those who were previously not able to afford them due to not being covered by any insurance scheme. As a result, providers have a greater population of potential patients. According to the theoretical literature, these healthcare providers may have an incentive to increase the price of their healthcare services to maximize profits to the greatest extent compatible with the continued financial viability for patients of being treated in their facilities. Feldstein (1970) showed that physicians raise their fees when insurance becomes more extensive—i.e., when a large part of the population comes to be covered by health insurance. Chiu (1997) uses a formal model to show that the introduction of insurance increases the equilibrium price of healthcare. In such cases, it is uncertain whether or not insurance will

reduce the financial burden of sick people. Chiu (1997) shows that if the supply of healthcare is sufficiently price-inelastic, this increase in price always leads to a reduction in consumer welfare.

### 3 Data and Descriptive Statistics

Our data combines the results of the longitudinal Rural Fixed-Point Survey (RFPS), carried out between 2003 and 2006, and those of a supplementary household survey aimed at evaluating the NCMS. The RFPS has been carried out each year since 1980s under the responsibility of the Chinese Ministry of Agriculture. Over the years 2003-2006, the same households were surveyed, with weekly book accounting recorded by these households as the primary information source. The sample was selected on the basis of a multi-stage stratified random sampling strategy. The 2006 round covers 26 provinces and 19,488 households. It provides information on individual demographics, including health status, and household characteristics, including income and expenditures.<sup>4</sup> The supplementary survey was conducted by Tsinghua University in May 2007. It surveyed a subsample of the 2006 round of the RFPS, collecting detailed information about the time when a county and a household enrolled in the NCMS.

The final sample excludes some outliers, such as households who have exited the NCMS or who participated in some cooperative insurance program between 1993 and 2002. Also excluded are households who purchased commercial insurance in 2007. The final sample includes 520 villages, 17,715 households and 71,866 individuals over the 4-year sample period.

Table 1 shows the participation rates of counties and households between 2003 and 2006. The percentage of our sample counties enrolled in the program increased over the years, from about 16% in 2003 to about 77% in 2006. Similarly, the enrollment of individuals also increased over the years, from 9% in 2003 to 72% in 2006. These numbers are consistent with the national data (Chinese Statistical Yearbook 2010). In the villages within counties which had launched the NCMS (called NCMS counties), the majority of households participated in the program, with the participation rate increasing from 64% in 2003 to 95% in 2006. Moreover, most households

---

<sup>4</sup>There are quite a few mistakes in the identification code used to track individuals and households. We used conservative rules based on individual's age, sex, and education to match individuals and households over years. If more than half of the household members could not be matched across two years, we excluded the household from our sample. Altogether, we deleted around 8% households from our sample due to the inconsistency of the identification code.



started to participate in the first year that their county launched the program—the participation rate by county in the year following the launch of the program was 64% in 2003 and 96% in 2006. Nevertheless, these numbers also indicate that quite a few households chose not to participate in the program or delayed in participation. Over the four years from 2003 to 2006, about 12.8% of households in NCMS counties did not participate.

Table 2 displays the descriptive statistics on self-reported health status. It is based on the answer to the question "How do you feel about your health status?" and there were five options: Excellent, Good, Fair, Bad, Very bad. As can be seen in the table, more than half of the sample felt excellent about their health, while about 12% of individuals reported fair or worse health. Over the years 2003-2006, the trend in the number of people who felt excellent was positive. Average SAH went from Good to Excellent. Table 2 also relates the subjective health status to enrollment status in the NCMS. The results highlight that both the insured and non-participants experienced an improvement in health status over the period studied. In fact, non-participants experienced even greater increases in subjective health than the insured. In contrast, the non-exposed experienced a deterioration in subjective health. However, this might result from dynamic selection, as it is possible that people who continue not to participate are those with excellent health, and that villages in counties which continue to be excluded from the NCMS program are those with bad average health.

Table 3, Column (4) gives preliminary statistics on individual, household and village variables on the whole sample (2003-2006). An inhabitant with no education is the reference. Women made up 47.5% of the sample, and the average age was 37. About 27% of the sample are doctors, and the villages on average deliver information on health more than 3 times a year. We also have information on chronic disease (such as Hypertension, Heart diseases, Nephritis, and Diabetes), other serious diseases (hepatitis A, hepatitis B, phthisis, AIDS, schistosomiasis), and activities of daily living (ADLs: cleaning yards, bowing and bending knees, bathing, feeding, dressing, getting out of bed) for each individual in 2007. Specifically, 7.5% of the sample suffer from a chronic disease. One half of the rural inhabitants in the sample have a primary school education or less. The proportion of laborers was around 55%. The percent of minority ethnic households was about 10% of the sample, which corresponds to the official statistics. Differences were observed between the samples: among non-participants and the non-exposed the percent of minority ethnic households was more than 15%. Average household size was around 4 members. On average,

households included at least one member below the age of 10 years, and at least one over the age of 65 years. In China, the size of the household is restrained by the one-child policy. While in rural areas, the policy constraint is less strictly enforced than in urban contexts, households still face monetary penalties according to the number of excess children. Moreover, children are quite often raised by their grandparents. For reasons related to finances, tradition, and registration permits (‘hukou’)<sup>5</sup>, the children remain in their hometown and may be a source of concern for their elderly caretakers.

*Wubao* individuals are those classified as under the poverty line as defined by the local government. They are then eligible for public policies aimed at helping the poor. They made up around 0.2% of our sample. Similarly, 18% of villages were considered *pingkun*, meaning that the village as a whole was under the poverty line as defined by the local government. In such cases, the whole village is eligible for governmental help, which can be financial or involve infrastructure improvements. Thirty-eight percent of villages were classified as *xiaokan* by the local government. These villages possess enough wealth to be above the poverty line defined by the local government.<sup>6</sup> Fifteen percent of the households lived in the suburbs and 14% were located in a “township”. 87% of the towns are in areas administratively defined as “agricultural regions”. In topographical terms, half of the villages are located in hill or mountain areas.

The data presented in Table 3, Columns (1) to (3) also shed some light on the differences between the insured, non-participants and non-exposed rural inhabitants. Since the number of counties and households participating in the program increased over time, 2006 includes more observations for the insured group, while for 2003 there are more for the non-exposed group. To ensure that the values of the variables, which may change over time, is comparable between these three groups, we focus on their values in 2003. Subscribing households generally had higher income and include fewer members who are older than 65 or younger than 10. The heads of these households were slightly older, more educated, less likely to be female, single, or a non-agricultural worker. They were also more likely to include members of the Communist Party, and less likely to be members of a minority or in poverty (*wubao*). Among non-participants and the non-exposed the percentage of ethnic minority households was more than 15%, but in the insured group it was under 10%. Villages in NCMS counties were also much richer and had fewer clinics but a higher

---

<sup>5</sup>Migrant workers’ children have very limited access to public schools in the regions where they work

<sup>6</sup>The criteria defined by individual local governments for classification as *wubao*, *xiaokan* and *pingkun* are set up at the local government level and they are highly variable.

rate of child vaccination than those in non-NCMS counties. They also included fewer migrants, more laborers, had higher average education levels and were less likely to be located in mountain, western or central areas.

## 4 Econometric Strategy

The subject of our modeling is the response to the survey question "How do you assess your health?", where answers range from very bad health (coded as 1) to excellent health (coded as 5) on a five-point scale.  $Y$  is measured on a discrete ordinal scale. We use an ordered probit model. Our empirical analyses exploit the quasi-natural experiment arising from the introduction of NCMS programs to examine their effect on the self-assessed health of individuals. As mentioned above, each period includes three types of households: participants in the NCMS program, non-participants within NCMS counties, and non-exposed households located in non-NCMS counties.

To learn how the introduction of the insurance program affected individual self-assessed health, we begin by applying the difference-in-difference (DID) framework. More specifically, the effects of the NCMS are identified through differences between insured and uninsured households in changes in SAH over the time periods before and after the local launch of the NCMS. The advantage of this double-difference framework is to eliminate time-invariant selection bias. This is crucial in the present context because participation in the program is voluntary, meaning that households which chose not to participate may differ from participants in both observable and unobservable characteristics. Moreover, program distribution over counties may be non-random. As a result, the members of insured and uninsured households may have had differing SAH even in the absence of the NCMS. The double-difference method can deliver unbiased and consistent estimates as long as the changes in household SAH over time would have been parallel were there no NCMS. Sufficient conditions are: (a) that trends in period-specific unobservables are the same in treatment and control groups, and (b) the expected change in idiosyncratic errors is zero for both the treated and untreated. Hence the claim that double differencing makes it possible to estimate the impact of a program if the unobservables that differ between the treated and the untreated are time-invariant (Wagstaff et al., 2009).

The introduction of the NCMS program may have had two types of effects: one, that of insurance coverage, has been widely studied in the health insurance literature. The other is the effect

of the introduction of a large-scale public insurance program, and other concomitant changes in healthcare provision, on both, participants and non-participants. Indeed, the introduction of the NCMS coincided with other relevant changes: in particular, governments implemented supporting policies to improve the quality and delivery of healthcare services. In addition, anecdotal evidence indicates that the price of healthcare services increased after the introduction of the NCMS (Mao, 2005). To identify the **insurance effect** of the program, which results only from the insurance coverage itself and excludes the effects of the aforementioned contemporary policies and changes, we use the double difference method to estimate the difference between participants and non-participants in NCMS counties, because both groups were affected by these changes. By contrast, the double difference between non-participants and the non-exposed yields what we call the **spillover effect** of the NCMS on non-participants. Refinements such as matching DID and tests of the identification assumptions are discussed in Section 6.

## 4.1 Difference-in-Difference Model

To facilitate understanding, we first present the model where the treatment is participation in the program given that the county has enrolled. The regression model for the double-difference comparison between participants and non-participants is as follows:

$$Y_{ihjt}^* = \beta X_{ihjt} + \gamma HHinsured_{ht} + \tau_t T_t + \delta HH_h + \epsilon_{ihjt}^1 \quad (1)$$

$$HHinsured_{ht} = 1 \text{ if the household } h \text{ participates to the NCMS program at the year } t \text{ such as } t > t_0 \quad (2)$$

where  $Y_{ihjt}^*$  represents the SAH of individual  $i$  of household  $h$  located in village  $j$  in period  $t$ . Here,  $t_0$  is the year just before the first year of the household's participation.  $HHinsured_{ht}$  is the binary variable that indicates whether household  $h$  has subscribed to the NCMS in year  $t$ .  $T_t$  includes three year dummies that control for the trend over time that is common to all groups. This assumption of a common trend for all groups is discussed in the Robustness section.  $HH_h$  includes all the household indicators.  $HH_h$  controls for all time-invariant household characteristics.  $X_{ihjt}$  includes observable individual variables that vary over time and may affect SAH (See Table 3 for details).  $\gamma$  represents insurance **coverage effects**, as it shows the difference in outcomes between participants and non-participants.

We now include the non-exposed.

$$Y_{ihjt}^* = \beta X_{ihjt} + \gamma HHinsured_{ht} + \tau_t T_t + \delta HH_h + \kappa CountyEnrolled_{jt} + \epsilon_{ihjt}^2 \quad (3)$$

$$CountyEnrolled_{jt} = 1 \text{ if the county } j \text{ enrolled in the NCMS program at the year } t \text{ such as } t > tt_0 \quad (4)$$

Here,  $tt_0$  is the year just before the year of the county's enrollment. The SUTVA assumption is discussed in the Results section.  $\kappa$  gives the **spillover effects** of the NCMS program showing the difference in outcome between the non-exposed and non-participants.

## 4.2 Results: NCMS Insurance Program

We now present the results on differences in dynamic changes in individual SAH. The difference-in-difference (DID) model applies a fixed-effect regression that controls for fixed effects of county, household and year. All the time-invariant effects of household characteristics are controlled by household fixed effects. In our context, the program is voluntary and households chose whether or not to participate. All the time-invariant effects of county characteristics are controlled by the county fixed effects. Similarly, enrollment in the NCMS program was not compulsory for counties. Counties chose whether or not to enroll. Note that households did not change county during our period of observation. Hence, there was colinearity between county fixed effects and household fixed effects. We thus only consider the household fixed effects that include the county fixed effects. On average (or median), household fixed effects are computed on 14 observations. The yearly time trend in SAH that is common to all individuals is controlled for with year fixed effects. The reference is a NCMS non-exposed man aged between 20 and 30 years old, non-educated, in 2003, living in a non-governmental township, in eastern and plains areas.

Controlling for the selection issue changes the results on the coverage effect and the spillover effect (Table 4, Columns (1) *versus* (2)). When no selection bias is considered, coverage appears to deteriorate subjective health, whereas the presence of general public health insurance leads to an improvement in SAH. Taking selection bias into account leads to a reversal of the coverage effect of the NCMS program: household NCMS subscription improves individual SAH. The spillover effect of the NCMS program does not significantly affect self-reported health. To sum up, after controlling for differences in county (household) characteristics according to the year of NCMS program introduction (NCMS household enrollment), a household's coverage by the NCMS improves perceived health.

The amount of time that has passed since the NCMS program was launched may impact its spillover effect on SAH. In the first year following its launch, the NCMS may have had a lesser impact than the following years. Several explanations may be advanced for this observation: as the years pass, increasing interest in health, increasing trust in the NCMS program,<sup>7</sup> as well as improvement in the community’s feelings of satisfaction with respect to public provision of healthcare and healthcare access (by central and local governments) through the NCMS program. We then allow for the possibility that the average treatment effect varies over the course of the implementation cycle of the NCMS program by considering an additional effect: number of years following the first year of county NCMS enrollment. Columns (3) and (5) give the results of this analysis. There is a highly significant positive association between seniority of NCMS county enrollment and self-assessed health. This additional effect dominates the overall spillover effect of the NCMS program.

Similarly, the decision to subscribe may affect the coverage effect on SAH. To assess how this effect varies over time, we also consider an additional variable, distinguishing the years after a household subscribes from the NCMS subscription by itself. This seniority effect is comparable to an intensive margin effect. Columns (4) and (5) give the results of this analysis. The seniority of household NCMS subscription (more than one year) has no effect on the SAH. Therefore, what really matters is the fact of subscribing, and not the length of subscription. In other words, there is an extensive margin but no intensive margin.

The model now controls for both the seniority of the launch of the NCMS program and the seniority of the NCMS household subscription (Column (6)). The seniority of household NCMS subscription (more than one year) still has no significant additional effect beyond the overall coverage effect. Therefore, the length of a household’s coverage by the NCMS program does not improve self-reported health. One explanation is the high correlation between the coverage effect and county enrollment after one year. Indeed, as shown in Table 1, from 2004, when a county had enrolled in the NCMS program, the percentage of subscribing households jumped to around 90%. The spillover effect increases with the seniority of the county’s enrollment in the NCMS program. Hence, the earlier the county enrolled in the NCMS program, the stronger the effect.

---

<sup>7</sup>Indeed, different local health insurance programs had previously been launched, with poor results in terms of improving healthcare access. As a result, public health insurance programs were widely considered to be untrustworthy.

## 5 Robustness

The results of the robustness tests are presented in Table 5.

To test whether the identification assumption of the baseline model holds, we apply the same model to alternative samples. The primary concern with the baseline model is that households that enrolled in the NCMS may not be comparable to households that did not. Particularly, we are aware that households that had not joined by 2007<sup>8</sup> may have already been covered by other insurance policies or had very specific concerns regarding participation. Therefore, in sample B, we excluded counties that had not participated by 2007. Given that the program was first piloted in July 2003, we exclude observations for 2003 from counties that launched the program in 2003, in order to avoid complications resulting from the start of the NCMS in the middle of that year. As a result, the year 2003 can be treated as the year where no counties introduced the NCMS: sample C. However, the counties that enrolled in 2003 were selected by the central government on economic and social criteria. The year 2003 was an introductory period for this reform. These counties may have some specific characteristics that are not well captured by the DID method. In sample D, data from 2003-enrolled counties are eliminated for all years.

In this paper, we use an ordered probit model with fixed effects. As mentioned above, household fixed effects are computed on an average/median of 14 observations. However, for some households, the data includes only 2-5 observations. Therefore, we may have an incidental parameters problem. One robustness check is to run the same model using only data from households with more than 14 observations. We call this sample E.

Analyses with different samples (Sample B, Sample C, Sample D, Sample E) confirmed the extensive margin effect. In all samples, the intensive margin has no effect on subjective health. The other factor whose effect is consistent for all the samples, is the spillover effect for villages in NCMS counties with seniority of enrollment.

In the literature on the effects of NCMS, authors such as Wagstaff (2009) consider the adverse selection issue—i.e., the possibility that insured individuals have specific characteristics that differ from those of non-subscribers. Table A.1 in the appendix gives the regression of household participation decisions against observable household characteristics. This analysis confirmed that

---

<sup>8</sup>The data also contains information on subscriptions for the year 2007

most of the differences in household characteristics are significant. Table A.1 also displays the regression of the county enrollment. Part of the differences in county and village characteristics are significant. Moreover, the evidence for adverse selection is mixed: the participation rate increases with the share of members with good or fair health status, when the reference group is the share of members with excellent health status; however, participation rate actually declines when the share of members with bad health status rises.

We couple the DID approach with matching to reduce the selection bias on the observables. The matching method can address two types of potential biases in the baseline method: bias caused by difference in the supports of the observable covariates between treated and untreated groups, and bias due to the difference between the two groups in the distribution of the related observables over their common support (Smith and Todd, 2005). The matching method applied here is propensity score matching. We estimate the propensity score in two steps. First, in each year, we estimate the probability of household enrollment in the villages where the county has launched the program. This is estimated on the basis of households in the NCMS counties. We predict the probability of enrollment of households in non-NCMS counties on the basis of this estimate. Second, we estimate the probability of county enrollment in the NCMS in each year. For both steps, a probit model is estimated. For the comparison between the insured and non-participants, the probability predicted in the first step is the propensity score used in the matching. For the comparison between the insured and non-exposed, both the similarity between households and that between counties must be considered; hence, the (composite) propensity score in each year, i.e., the product of the two probabilities predicted from the aforementioned two steps, is used in the matching estimation. Figure A1 shows the histogram for the (composite) propensity scores for the three groups: the insured, non-participants, and non-exposed. As expected, the distribution of the propensity score is more skewed to the right for the insured than for non-participants. Nonetheless, the region of common support is adequate. The propensity score distribution for the non-exposed is more skewed to the left than the one for non-participants because of the lower possibility of county enrollment.

Given that we are considering more than two periods, the traditional DID matching method requires modification. We first estimate the insurance effect in each year from 2004 to 2006. Then we also determine an alternative measure: a weighted average of the insurance effect over these three years calculated by weighting the effect in each year on the basis of the ratio of the number



of treated in that year to the total number of treated over the three years. The standard errors are bootstrapped with 100 replications. For each year, we use five-nearest-neighbor matching with replacement and caliper 0.01, and impose a common support condition. The distance between the propensity scores is measured by the Mahalanobis metric. We confirm that the results are not sensitive to the number of neighbors (including 3 and 10) or the choice of caliper (including 0.005 and 0.0025). Another complication here is whether to consider experienced participants (those who are in their second or later year of enrollment in the NCMS) among the treated. To test the robustness of our results, we also study the case where all the experienced participants are excluded. In this case, we focus on the first-year effect of the NCMS on the SAH.

To summarize, *Case 1*: On the whole sample, probabilities are computed year by year. *Case 2*: On the whole sample, a weighted average of the insurance effect is computed. *Case 3*: Excluding all the experienced participants, probabilities are computed year by year. *Case 4*: Excluding all the experienced participants, a weighted average of the insurance effect is computed.

In all cases, the result with the DID matching model is the same. NCMS household subscriptions improve individual subjective health but seniority has no effect. A spillover effect is also observed but this effect depends on the duration of county enrollment: the effect of the NCMS program increases with the time that the county has been enrolled in the NCMS.

Furthermore, we assume a common trend for all groups, attributing any difference in trends between the treatment and control groups that co-occur with the NCMS program to the program. This assumption may be debatable, so we attempt to check its robustness on the results by performing a placebo test using a "fake" treatment group. Here, we present the results for the year 2005. The "fake" treatment group is the 2006-NCMS-households (i.e. households who had not yet subscribed in 2005) and the control group is the group of households without NCMS in 2006 (and thus also without NCMS in 2005). On this analysis there is no significant effect of the NCMS program. Hence, our original DID is not biased (results are available on request).

We also make the stable unit treatment value assumption (SUTVA), which implies that *i*) the treatment status of any county does not affect the potential outcomes of another county (non-interference); *ii*) the treatments for all counties are comparable (no variation in treatment). The supply side may be thought to change in non-treated areas because of the attraction of suppliers to NCMS-counties due to the increase in affordability of the healthcare for the NCMS-insured.

Thus, as a robustness check, we control for supply side variables. The results are shown in Table 4, last column. We then compare the NCMS effect with and without this control (Columns (5) and (6)). The absence of change in the NCMS effects suggests that the SUTVA assumption is valid.

## 6 Individual Characteristics and Geographical Effects

This section presents an overview of determinants of individual subjective health. Readers who are interested only in NCMS program effects may skip this section. Results are presented in Table 6. First, as expected, age has a negative effect on health index.<sup>9</sup> Women’s self-assessed health tends to be poorer than that of men. Being single also decreases health status. This could be due to greater feelings of insecurity or due to taking less care of their health. The survey includes questions regarding individual physical health and individual health behavior. SAH is worse when the individual suffers from a chronic disease or common serious illnesses (hepatitis A, hepatitis B, phthisis, AIDS, schistosomiasis).<sup>10</sup> We tested the effects of variables such as social class and labor market status to capture individuals’ economic and financial status. Working in the industrial sector or having a non-agricultural job significantly improves SAH. The model reveals that education and income are strongly correlated with self-assessed health. The higher an individual’s level of education, the better their health. This correlation partially results from an ”age cross education” effect, with older respondents being less educated on average due to increasing access to secondary and higher education over the past thirty years. Higher household income is associated with better perceived health.<sup>11</sup>

Because of colinearity problems, focusing on the effects of the geographical variables and household variables requires to regress the model without county fixed effects and household fixed effects. The results are displayed in Table 6.

Perceived health deteriorates with increasing household size. This may be explained by family composition. As explained above, the children of internal migrants remain in their hometown,

---

<sup>9</sup>As is well-known, in China there is massive internal migration of healthy 15- to 45-year-olds from rural to urban areas. Because of missing healthy people aged between 15 and 45 years, a polynomial specification for age is not relevant here. The results of the model with polynomial specification for age are available on request.

<sup>10</sup>The data does not provide detailed information on these pathologies.

<sup>11</sup>Here we do not test the income-related heterogeneity in self-assessed health.

i.e. their village, with their grandparents. This may also be a cause of worry or fatigue in elderly grandparents.<sup>12</sup> SAH worsens with the number of old people (over 65 years old) in the family and with the number of young people (less than 10 years old) in the family.

Members of ethnic minorities are more likely to report being in bad health. Even after controlling for the level of village income and village status (as being under the poverty line as defined by the province), minority status remains significantly associated with worse health status. Being a very poor family as defined by the local government (*wubao* household) has no significant effect after controlling for household income and for being in a *pingkun* village (or not). A village classified as *pingkun* by the local government receives some subsidies. Higher household income is associated with better perceived health.

Moreover, having a doctor in the household or a household member with training in primary care is not associated with better perceived health. In contrast, having a cadre of the Communist Party in the household makes individuals more optimistic on their own health. A household member having a party position might help other household members to get a job and to resolve certain issues involving local government.

The estimated coefficient of living in an agricultural region on reported health is positive, whereas living in a “townships” deteriorates self-assessed health. Township areas offer more healthcare facilities and/or more access to healthcare facilities, but the price of healthcare is also higher than in the countryside. Furthermore, some other elements whose aim is to improve healthcare access are associated with this program. These include the expansion of healthcare facilities, free health examinations (in general, one per year), and medical training education. The presence and extent of these elements depends on decisions of the local government, and their effects are unclear. While vaccinations for children improve individual health, the effect of free clinical examinations and health training education are more ambiguous. Having access to more information on health status may also make people more aware of their health problems.

## 7 Conclusion

During the 2003-2006 period, subjective health status in Chinese rural areas improved. We use a data for these years from a household longitudinal survey to analyze how the introduction of

---

<sup>12</sup>The children of internal migrant workers have very limited access to public schools in the areas where their parents work.

an insurance system, known as New Cooperative Medical Scheme (NCMS), contributed to this change. The health insurance program was launched sequentially over counties beginning in 2003, and households voluntarily choose whether participate in the program after it has been launched by their county. These facts make it feasible to disentangle two channels of influence of the NCMS on self-reported health status: one is the insurance effect of the coverage, which can be estimated by comparing insured households with non-participants in counties that have launched the program (NCMS counties); the other is the general equilibrium effect that affects all residents of NCMS counties, which can be estimated by comparing non-participants to non-exposed households (located in counties where the program has not been implemented). The longitudinal data also allow us to examine how the effect changes with the duration of households' subscription to the program. In this paper, the variable of interest is the subjective health in itself. We do not consider subjective health as a proxy of objective health. We focus here on perceived individual health, a determinant of demand for healthcare access and treatment.

The empirical findings include, first, a positive extensive margin: individuals feel better about their health status when they are covered by the NCMS. However, there is no intensive margin: there is no additional gain in self-assess of health status when the individual enrolls in the program for more years. Second, we find a positive general equilibrium effect of introducing the NCMS program on non-participants in the NCMS county, which accumulates over time.

Third, besides the insurance program, household composition (family size, number of young and old members) and socio-economic status (income, occupation, and education) also have strong effects on the self-evaluation of health status. Not only is larger household size associated with worse self-assessed health, but the composition of the household also has an effect: a larger number of old people is associated with worse perceived health, while the number of young people has a significantly positive association with perceived health. Both education and income have an effect on SAH. A high level of education and a relatively high income lead to a better perceived health. Controlling for wealth and geographical area, it turns out that ethnicity still has an impact: Being part of a minority has a significant negative effect on self-reported health. Having a position in the Communist Party has a positive effect on subjective health.

Our findings focus on the role of the NCMS program in improvement in SAH in Chinese rural areas. Understanding the effect of the implementation of a health insurance system on self-reported health may help policy makers to design insurance schemes with increased participation

rates and ensure that rural populations gain better access to health services and enjoy greater protection against the cost of illness.

## 8 References

- Bai, C. and Wu, B. (2011), "Health Insurance and Consumption: Evidence from China's New Cooperative Medical Scheme", Working paper series
- Bloom, G. and Gu, X. (1997), "Health Sector Reform: Lessons from China", *Social Science and Medicine*, 45, 351-360.
- Chiu, W. (1997), "Health insurance and the welfare of healthcare consumers", *Journal of Public Economics*, 64(1), 125-133.
- Feldstein, M. (1970), "The Rising Price of Physicians' Service", *Review of Economics and Statistics* 52, 121-133.
- Gao, J., Tang, S., Tolhurst, R. and Rao, K. (2001), "Changing access to health services in urban China: implications for equity", *Health Policy and Planning*, 16, 302-12.
- Hesketh, T. and Zhu W. (1997), "Health in China: The healthcare market", *British Medical Journal*, 314, 1616-8.
- Hung, E. and Chiu, S. (2003), "The Lost Generation: Life Course Dynamics and Xiangang in China", *Modern China*, 29(2).
- Lei, X. and S. Lin (2009), "The New Cooperative Medical Scheme in Rural China: Does more coverage mean more service and better health?", *Health Economics* 18, 25-46.
- Liu, Y. (2006), "What is wrong with China's health system", *Harvard China Review* 3:14-18.
- Milcent, C. (2012) "Industrialisation and Inequality: Healthcare Use in Rural China", Working Paper, PSE series.
- Koszegi, B. (2006), "Emotional agency", *The Quarterly Journal of Economics*, 121(1), 121-155
- Smith, A. and Todd, P. (2005), "Does matching overcome LaLonde's critique of nonexperimental estimators?", *Journal of Econometrics*, 125(1-2), 305-353.
- Wagstaff, A., M. Lindelow, J. Gao, L. Xu and J. Qian (2009), "Extending health insurance to the rural population: An impact evaluation of China's new cooperative medical scheme", *Journal of Health Economics* 28, 1-19.
- Wagstaff, A. and S. Yu (2007), "Do health sector reforms have their intended impacts? The World Bank's Health VIII Project in Gansu province, China", *Journal of Health Economics* 26 (3): 505-535

**Table 1: Participation rate in different years**

Year	2003	2004	2005	2006	2007 or later
<b>County participation</b>					
Cumulative enrollment rate	16.4%	32.1%	45.7%	77.1%	100%
<b>Household participation</b>					
Cumulative enrollment rate	9.5%	23.8%	40.0%	72.3%	100%
Participation rate in NCMS counties	63.7%	79.8%	87.0%	94.6%	
Exposure rate	83.2%	66.8%	51.2%	22.0%	
Cumulative participation rate in counties that launched the NCMS in 2003	63.7%	82.1%	94.4%	97.4%	100%
Cumulative participation rate in counties that launched the NCMS in 2004		75.8%	93.9%	95.0%	100%
Cumulative participation rate in counties that launched the NCMS in 2005			71.2%	87.0%	100%
Cumulative participation rate in counties that launched the NCMS in 2006				96.2%	100%

**Table 2: Self-reported Health Status**

	All years	2003	2004	2005	2006
<b>Part I</b>					
Health category					
Very bad: 1 credits	1.84	1.71	1.85	1.93	1.84
Bad: 2 credits	3.03	2.96	2.98	3.08	3.11
Fair: 3 credits	7.55	7.91	7.79	7.25	7.27
Good: 4 credits	34.43	36.64	34.58	33.78	32.93
Excellent: 5 credits	53.15	50.78	52.8	53.96	54.85
Mean value	4.34	4.32	4.33	4.35	4.36
Number of observations	71866	16892	18099	18301	18574
<b>Part II</b>					
Mean value of self-reported health status					
Insured	4.37	4.32	4.33	4.36	4.38
Non-participants	4.39	4.25	4.27	4.42	4.48
Non-exposed	4.33	4.34	4.34	4.33	4.27
Proportion of individuals with fair or worse health					
Insured	12.1%	13.6%	13.5%	11.8%	11.7%
Non-participants	11.5%	13.7%	13.6%	10.5%	10.2%
Non-exposed	12.5%	11.9%	12.2%	12.8%	14.2%

**Table 3: Descriptive Statistics for Three Groups**

Variable	Villages within NCMS counties		Villages within non-NCMS counties	Whole sample
	Insured	Non- participants	Non-exposed	
	<i>Based on year 2003</i>			<i>Years 2003-2006</i>
	(1)	(2)	(3)	(4)
Woman	47.8%	46.5%	47.4%	47.5%
Age	37.6	37.3	36.5	36.9
Serious disease	0.6%	0.5%	0.6%	0.6%
Chronic disease	7.2%	4.6%	7.8%	7.5%
Household size	4.6	4.6	4.6	4.6
Percent of members older than 65	33.0%	40.0%	32.0%	32.7%
Percent of members younger than 10	47.0%	48.0%	49.0%	48.1%
Age of head of household	51.7	50.2	50.5	50.9
Head's years of education	6.7	6.5	6.5	6.6
Female head	4.9%	6.2%	8.5%	7.1%
Single head	7.1%	10.5%	9.1%	33.9%
Head in industry sector	36.6%	42.4%	33.1%	34.8%
Having communist members	17.0%	10.0%	15.0%	15.7%
Having communist cadres	9.0%	6.6%	8.9%	8.8%
Including doctors or individuals with medical training	21.1%	19.6%	29.5%	26.1%
Minority household	8.7%	17.1%	15.2%	13.0%



<i>Wubao</i> household	0.2%	0.3%	0.2%	0.2%
Village average income per capita in 2003	4014.6	3812.2	3083.0	3446.1
Township	15.1%	10.7%	13.9%	14.2%
Surburb	14.1%	7.7%	14.0%	13.8%
<i>Xiaokan</i>	36.3%	20.6%	37.0%	36.0%
<i>Pingkun</i>	7.8%	5.5%	9.5%	8.7%
Number of clinics in 2003	129.0%	124.0%	129.0%	132.6%
Deliver information on health more than 3 times a year	60.1%	72.7%	57.4%	59.1%
Share of children vaccinated in 2003	97.2	98.7	97.4	97.1
Share of migrants in 2003	23.2%	22.1%	25.2%	24.3%
Share of laborers in the village	57.0%	54.5%	53.6%	54.9%
Share of laborers with primary school level or below	46.3%	51.3%	48.8%	48.1%
Share of laborers with high school degree or above in the village	33.8%	31.1%	30.2%	31.5%
Mountain area	47.4%	50.6%	57.8%	53.7%
Highland area	24.5%	18.4%	23.2%	23.5%
Western area	22.4%	20.9%	29.3%	26.5%
Central area	43.3%	35.6%	47.7%	45.6%

**Table 4: Ordered Probit Model with Fixed Effects (Dependent variable: SAH)**

	(1) <sup>@</sup>	(2)	(3)	(4)	(5)	(6)
Woman	-0.094*** (0.010)	-0.218*** (0.013)	-0.218*** (0.013)	-0.218*** (0.013)	-0.218*** (0.013)	-0.218*** (0.013)
Age	-0.032*** (0.000)	-0.053*** (0.001)	-0.053*** (0.001)	-0.053*** (0.001)	-0.053*** (0.001)	-0.053*** (0.001)
Single	-0.339*** (0.013)	-0.482*** (0.021)	-0.481*** (0.021)	-0.484*** (0.021)	-0.482*** (0.021)	-0.482*** (0.021)
Serious disease	-0.414*** (0.075)	-0.319** (0.131)	-0.320** (0.131)	-0.319** (0.131)	-0.320** (0.131)	-0.321** (0.132)
Chronic disease	-0.462*** (0.105)	-0.261 (0.161)	-0.260 (0.161)	-0.261 (0.161)	-0.261 (0.161)	-0.261 (0.161)
Edu. 3 to 6 yrs	0.314*** (0.015)	0.427*** (0.023)	0.426*** (0.023)	0.426*** (0.023)	0.426*** (0.023)	0.426*** (0.023)
Edu. 7 to 9 yrs	0.453*** (0.016)	0.623*** (0.025)	0.622*** (0.025)	0.622*** (0.025)	0.621*** (0.025)	0.621*** (0.025)
Edu: more than 9	0.577*** (0.022)	0.772*** (0.033)	0.769*** (0.033)	0.769*** (0.033)	0.768*** (0.033)	0.768*** (0.033)
Non agri. Self-employed	0.143*** (0.026)	-0.068* (0.036)	-0.071* (0.036)	-0.069* (0.036)	-0.070* (0.036)	-0.071* (0.036)
Non agr. Employees	0.046** (0.024)	0.115*** (0.034)	0.113*** (0.034)	0.114*** (0.034)	0.114*** (0.034)	0.114*** (0.034)
Non agri. other	0.026 (0.023)	-0.152*** (0.034)	-0.152*** (0.034)	-0.153*** (0.034)	-0.152*** (0.034)	-0.153*** (0.034)
Work in industry	0.065*** (0.021)	0.164*** (0.030)	0.165*** (0.030)	0.165*** (0.030)	0.165*** (0.030)	0.166*** (0.030)
NCMS HH subscription	-0.107*** (0.025)	0.116*** (0.042)	0.098** (0.042)	0.147*** (0.025)	0.095** (0.045)	0.065** (0.033)
Seniority of NCMS-HH subscription			0.066 (0.051)		0.023 (0.013)	0.025 (0.041)
NCMS County enrollment	0.177*** (0.024)	0.051 (0.041)	0.058 (0.042)	0.038 (0.041)	0.041 (0.042)	0.058 (0.042)
Seniority of NCMS-county enrollment				0.182*** (0.026)	0.163*** (0.043)	0.162*** (0.043)
No. Clinics						0.004 (0.015)
Child vaccins						0.002*** (0.001)
Medical exams						-0.002** (0.001)
Free medical prevention training						0.003 (0.005)
Year Fixed Effects	No	Yes	Yes	Yes	Yes	Yes
Household Fixed Effects	No	Yes	Yes	Yes	Yes	Yes

Standard errors are in parentheses. \* Significant at 10%; \*\*significant at 5%; \*\*\* significant at 5%.

@: No control for the selection issue (no fixed effects (years and households)).

**Table 5: Test of robustness**

	NCMS household subscription		Seniority of NCMS HH subscription		NCMS county enrollment		Seniority of NCMS county enrollment	
DinD Model								
Sample B	0.147***	(0.036)	0.023	(0.028)	0.039	(0.028)	0.159***	(0.026)
	0.97**	(0.034)			0.036*	(0.019)	0.187***	(0.027)
Sample C	0.133***	(0.035)	0.024	(0.027)	0.044	(0.042)	0.146***	(0.034)
	0.088***	(0.043)			0.041	(0.044)	0.146***	(0.035)
Sample D	0.123***	(0.049)	0.039	(0.028)	0.022	(0.038)	0.133***	(0.035)
	0.090***	(0.041)			0.018	(0.045)	0.153***	(0.035)
Sample E	0.102***	(0.036)	0.054	(0.044)	0.032	(0.038)	0.125***	(0.021)
	0.128***	(0.036)			0.053	(0.036)	0.087***	(0.038)
DinD Matching Model								
Case 1	0.084***	(0.046)	0.006***	(0.000)	0.015	(0.068)	0.067*	(0.041)
	0.089***	(0.046)			0.019	(0.068)	0.040*	(0.021)
Case 2	0.076	(0.045)	-0.002*	(0.001)	0.003***	(0.001)	0.104***	(0.036)
	0.060*	(0.028)			0.009	(0.007)	0.068***	(0.020)
Case 3	0.067**	(0.031)	0.051	(0.059)	0.028	(0.045)	0.134***	(0.001)
	0.071**	(0.037)			0.044	(0.042)	0.133***	(0.024)
Case 4	0.084**	(0.005)	0.040	(0.055)	0.029	(0.027)	0.108***	(0.000)
	0.102*	(0.054)			0.027	(0.030)	0.095***	(0.021)

Standard errors are in parentheses. \* significant at 10% ; \*\* significant at 5% ; \*\*\* significant at 5%. Estimation results on control variables are not reported in this table.

Regressions based on the set of variables presented Table 4, Columns(4) and (5).

**Table 6: Determinants of SAH (Dependent variable :SAH\_)**

Woman	-0.114*** (0.013)	Industry sector	0.047 (0.037)	Suburb	-0.223*** (0.039)
Age	-0.032*** (0.001)	HH size	-0.002 (0.012)	Live in township	0.139*** (0.039)
Single	-0.330*** (0.025)	No. of clinics	0.050*** (0.014)	Agricultural area	0.096*** (0.037)
Serious disease	-0.321*** (0.097)	Child vaccins	0.006*** (0.001)	HH member: w/communist card	-0.016 (0.038)
Chronic disease	-0.382*** (0.037)	Free exams	-0.001 (0.001)	HH member: w/communist party cadre	0.055** (0.026)
Edu. 3 to 6 yrs	0.253*** (0.027)	Health edu. training	-0.005 (0.004)	HH member: w/medical knowledge	0.107 (0.088)
Edu. 7 to 9 yrs	0.381*** (0.029)	HH Log(Income)	0.109*** (0.020)	HH Minority	-0.122*** (0.039)
Edu: more than 9	0.464*** (0.040)	Xiaokan village	-0.060** (0.029)	No. young in HH	-0.052** (0.023)
Non agri. Self-employed	0.109** (0.046)	Pingkun village	0.064 (0.049)	No. old in HH	-0.070*** (0.022)
Non agr. Employees	0.061 (0.042)	Log(village income)	0.237*** (0.031)		
Non agri. other	-0.020 (0.041)				

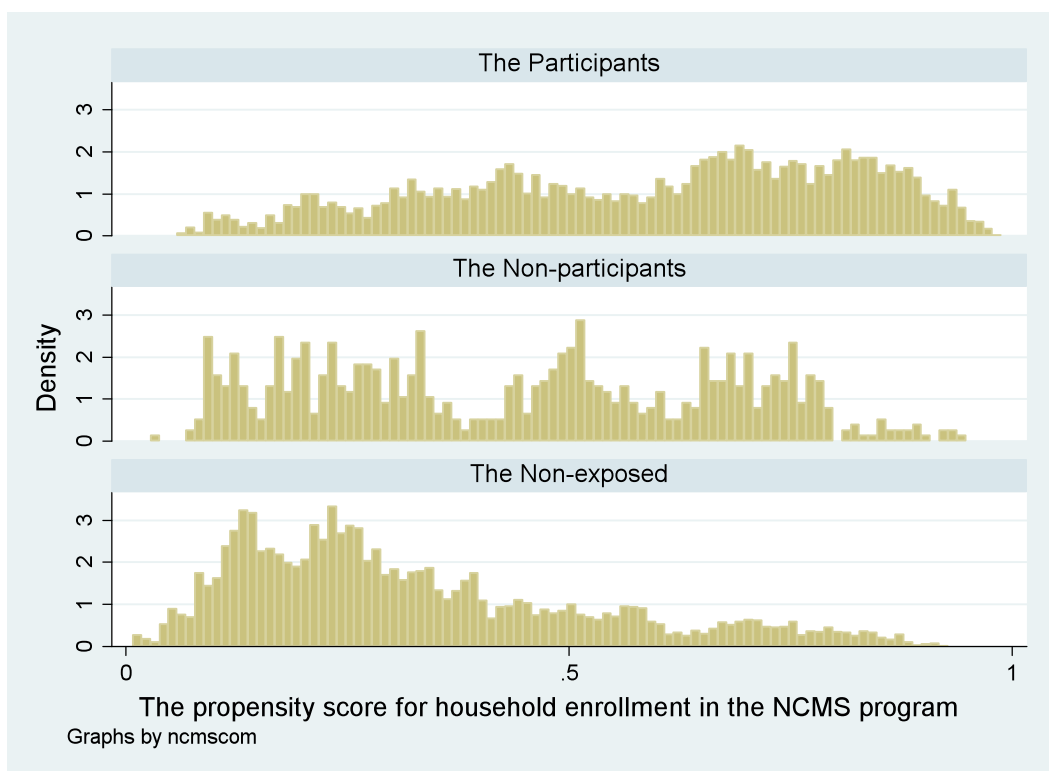
Standard error are in parentheses. \* significant at 10% ; \*\*significant at 5% ; \*\*\* significant at 5%.

#### *The specificity of the age variable*

Under the age of 15 years, health self-assessment may be debatable. Individuals may be too young to self-assess themselves. Dropping these observations does not change the results. We also regress the model with age divided into groups. For individuals born in 1958-1983, we divide age into 10-year categories, except over age 40 where we divide into 5-year categories. Considering age group dummies does not significantly change the results on the other independent variables. We find that members of the lost generation are more optimistic.

The lost generation is defined as the birth cohorts born in 1948-1957. Most members of this generation experienced disruptive education and send-down policy (Hung and Chiu, 2003). Using 4-year panel data from the longitudinal National Fixed-point Survey (NFS) 2003-2006, we defined 4 groups of age following Huan (2009). The members of the oldest cohort (Cohort 1, born in 1943-1947) came of age as young children and teenagers during the New China (1949 People's Republic China) and they either finished their high school education, college education, or entered the urban labor force before the beginning of the 1966-1976 Cultural Revolution. Cohort 2 (1948-1952) members experienced the early stages of the Cultural Revolution, in which send-down policy was applied seriously. They graduated from junior or senior middle schools in 1966, 1967, and 1968. Many of them spent 5-10 years in intensive labor on farmland before they were allowed to return to the cities. Cohort 3 (1953-1957) members completed secondary education and spent a few years in the countryside during the late stages of the Cultural Revolution. Cohort 4 (1958-1962) members spent their childhood during the Cultural Revolution. Disruption of education and send-down policy were specific to Cohorts 2 and 3, which are called the lost generation.

**Figure A1: Distribution of propensity scores by group.**



**Table A.1**

Household participation decision and county enrollment.

	Household participation decision <sup>a</sup>			County enrollment	
	(1)	(2)	(3)		(4)
Log(Income)	0.149*** (0.040)	0.139*** (0.044)	0.049 (0.054)	Log(village income)	0.443** (0.186)
Household size	-0.002 (0.017)	0.017 (0.020)	0.027 (0.026)	Villages classified as <i>xiaokan</i>	-0.352** (0.176)
Share of members with good health in 2003 <sup>b</sup>	0.207*** (0.070)	0.190** (0.080)	-0.105 (0.094)	Villages classified as <i>pingkun</i>	-0.154 (0.289)
Share of members with fair or worse health in 2003	0.459*** (0.142)	0.531*** (0.159)	0.076 (0.186)	Suburban area	0.129 (0.281)
Share of members with <i>poor</i> health in 2003	-0.868*** (0.169)	-0.813*** (0.179)	-0.643*** (0.215)	Township	-0.254 (0.326)
Share of members older than 65	-0.015 (0.139)	0.006 (0.148)	-0.109 (0.171)	Agricultural villages	-0.308 (0.254)
Share of members younger than 10	0.453** (0.200)	0.548*** (0.207)	0.321 (0.274)	Log(population)	0.153 (0.145)
Share of migrants in 2003	0.114 (0.132)	0.101 (0.141)	-0.095 (0.154)	Share of laborers	2.663*** (0.766)
Minority	-0.673*** (0.081)	-0.812*** (0.106)	-0.019 (0.216)	Share of high school or above	0.064 (0.728)
<i>Wubao</i> household	-0.690** (0.304)	-0.826** (0.355)	-1.311*** (0.440)	Share of migrants in 2003	-0.228 (0.482)
Household with communist members	0.216*** (0.074)	0.119 (0.078)	0.150 (0.095)	Number of clinics in 2003	-0.072 (0.080)
Female head	0.089 (0.112)	0.257** (0.116)	0.156 (0.145)	Share of children vaccinated in 2003	0.013** (0.006)
Age of the head of household	0.013*** (0.003)	0.011*** (0.003)	0.009*** (0.003)	Mountain areas	-0.470*** (0.180)
Head is single	-0.233*** (0.084)	-0.233*** (0.087)	-0.221** (0.103)	Highland areas	0.535** (0.219)
Head's years of education 3–6 <sup>a</sup>	-0.024 (0.101)	0.070 (0.104)	0.223* (0.127)	Western China	-0.602** (0.256)
Head's years of education 7–9	0.047 (0.108)	0.166 (0.114)	0.239* (0.136)	Central China	-0.334* (0.203)
Head's years of education 10 and above	0.122 (0.143)	0.230 (0.152)	0.043 (0.175)		
Head is self-employed non- farmer <sup>b</sup>	-0.237*** (0.074)	-0.286*** (0.086)	-0.074 (0.106)		
Head is an employee	0.131* (0.074)	0.017 (0.079)	0.026 (0.093)		
Head works in other non- farm-related employment	0.020 (0.071)	-0.023 (0.077)	0.184* (0.106)		
Observations	6575	6492	3960	Observations	394
Likelihood	-1904	-1711	-1154	Likelihood	-212.5
Pseudo $R^2$	0.194	0.273	0.406	Pseudo $R^2$	0.222
Year	Yes	Yes	Yes		Yes
Village fixed effects	No	No	Yes		No
Province fixed effects	No	Yes	No		No
Control for Village characteristics as defined in Column (4)	Yes	Yes	No		-

Note: Standard errors in parentheses, \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ .

a.- The omitted category is illiteracy or years of education less than 3.

b.- The omitted category is farmer. The regression for village enrollment is based on county-level data.